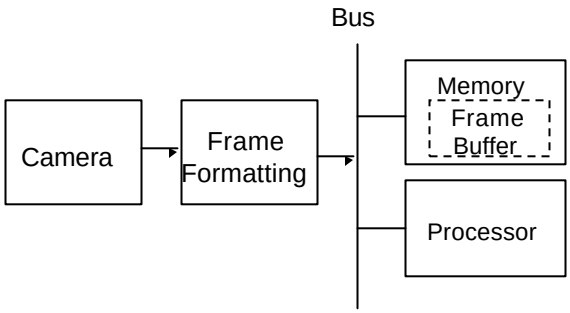
Networked Multimedia: Video



Lecture Overview

- Human vision - review
- Video standards
- Network video
- Video software

Color in Human Vision

review

- Human eye has two types of sensors
  - rods: monochrome
    - high resolution
  - cones: respond to one of three colors (RGB)
    - lower resolution
    - other colors in light spectrum seen because they stimulate two colors at once (*e.g.* cyan = G & B)
    - so we can display all colors by adding RGB
  - color component can be less precise than the monochrome image - not noticed
    - color TV takes advantage of this; less bandwidth devoted to color
    - in general computer displays color every pixel

Resolution of Human Vision

review

- Visual persistence causes impression of constant or moving image if frame rate is above 25 Hz
  - but psychological fatigue is experienced up to 60 Hz
  - some displays run at 72 Hz and above to avoid this
- Center of vision field has much greater acuity
  - can be used to reduce data rate elsewhere in the scene
  - if you can be sure the audience will look where you want them to...

Perceptions in Vision

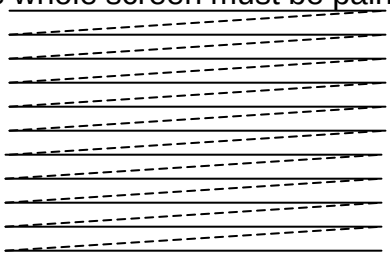
review

In general:

- Pixels blend into smooth image at the distance where they are too small to see individually
- Moving image perceived as better quality than still image of the same resolution
- Color images perceived as better quality than the same resolution monochrome
- Horizontal and vertical perceived differently
  - No stereo, and limited range, in vertical direction
  - Immersive effect greater if image occupies most or all peripheral vision
  - Thus CinemaScope movies and HDTV use wide wide aspect ratio (5:9, 9:16, or greater versus regular 3:5)

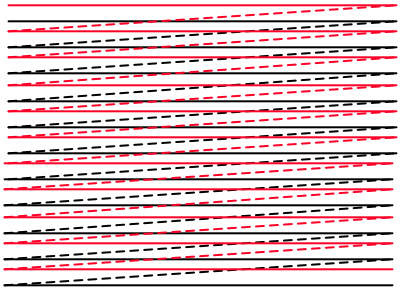
Raster Displays

- At one time there were two contending technologies for graphic display: in addition to raster (see below) vector graphics involved moving a beam around the screen in arbitrary patterns
  - good for drawing curves - like an oscilloscope
- Raster won, thanks to video RAM
- By contrast, video has always used a raster
  - because the whole screen must be painted every frame



Interlaced Raster Displays

- The National Television System Committee (NTSC) standard formalized in 1953 for US color TV faced a problem in that consumer technology did not have performance for a flicker-free 60 Hz scan at the minimum acceptable line rate (525 lines per frame)
- So they settled on scanning out half of the lines per field, interlaced:



US/Canada Broadcast Video Standard

- NTSC
  - compatible with previous US monochrome
  - 6 MHz channel
  - interlaced (generates visual artifacts)
  - chrominance (color) encoded at lower bandwidth than luminance
  - red encoded as difference between monochrome luminance and other colors
    - – can't display a really saturated red color
  - deployed ten years before European systems
    - – but then took ten years to catch on

European Broadcast Video Standards

- Developed after technology was more mature
- Sequential Couleur Avec Memoire (SECAM)
  - France and Russia
- Phase Alternating Line (PAL)
  - everybody else
- 25 Hz progressive scan
- 7 or 8 MHz channels
  - allows more lines (625)
- Encodes red separately
  - better saturated reds than NTSC

HDTV Standards

- Analog (all 16:9 aspect ratio)
  - Europe: 1250 lines
  - Japan: 1125 lines
  - US (proposed): 1050 lines
- Digital
  - self-defining formats allow receivers to select among a range of lines and aspect ratios
  - examples: 1080i, 720p, 480i, 480p
  - top-end: 1125 (1080 usable) by 1920 pixels
  - digital broadcast offers much clearer picture
    - but no fringe reception
  - FCC aiming for full urban deployment by 2006
    - simultaneous NTSC delivery until at least then

Data Compression: Differential Encoding

- Basic idea: transmit only the significant changes from some previously sent information.
- For example, the Discrete-Cosine Transform (DCT) algorithm used to send differences in color and intensity between bit-mapped image frames.
  - related to Fourier Series presented previously
  - some details lost in reconstituting
- Other examples:
  - Differential Pulse Code Modulation (see previous slides)
  - Motion prediction algorithms used for MPEG encoding
- Works on the principle that not many significant changes occur between the base image and the one(s) to be encoded; thus, most matrix cells will be zero

Multimedia Compression

- JPEG recodes still images using DCT
  - then drops information corresponding to very high frequencies
- MPEG recodes motion images to take advantage of redundancy between successive frames
- MP3 audio compression (developed for use with MPEG video) dynamically adapts its quantization to the program material
  - achieves 12:1 compression
  - used by controversial “Napster” system to share music recordings over the Internet
  - court killed Napster but others have taken its place

Video Compression

- Digital TV and video would be impractical without compression due to extremely high data rate
- Some formats:
  - Microsoft .avi - video, audio and labeling
  - Apple QuickTime .mov - comprehensive multimedia format for compressed video, audio, and graphics
  - MPEG - multiple generations
    - MPEG-1 computationally intensive to encode
      - decoding by consumer devices practical
    - MPEG-2 for DVD and HDTV achieves 55:1 real-time
- Multipurpose Internet Mail Extensions (MIME) use MPEG, JPEG

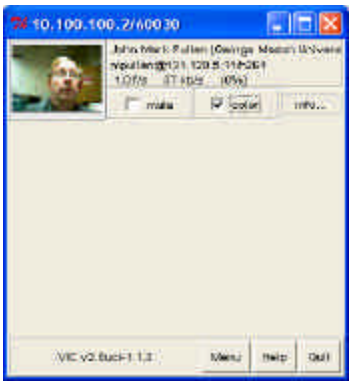
New Web Multimedia

- "Streaming" audio and video
- Custom clients play audio and video
  - examples: RealAudio, RealVideo
- Use TCP like other Web browser processes
  - use enough compression to fit in selected channel
  - for RealAudio and RealVideo, a 28.8 kbps modem
- Download and buffer enough data to give high confidence Internet "stalls" will not disrupt playback
  - this precludes real-time use but enables operation without provisions for QoS
- Play the stream as the rest of the file is downloading
  - use TCP flow control to avoid client buffer overrun

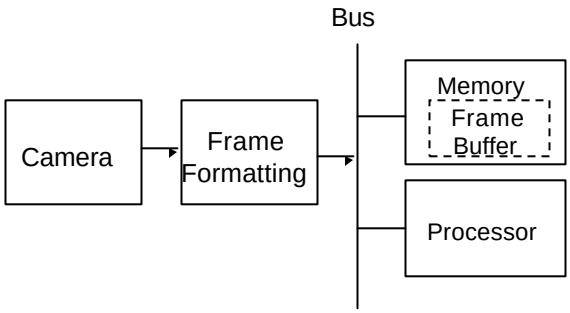
MBone Video

- Internet Multicast Backbone (MBone) stimulated much useful work in synchronous multimedia
  - including the VIdeo Conferencing tool (VIC)
  - developed by Lawrence Berkeley Lab (LBL)
  - expanded to Windows by University College London (UCL)
  - data rates 1 kb/s to 3Mb/s
  - frame rates 1 frame/s to 30 frames/s
  - encodings include jpeg and ITU-T h261, h263
  - display Common Intermediate Format (CIF)
  - sizes QCIF 160x120, CIF 320x240, SCIF 640x480 plus comparable PAL and NTSC

VIC Interface



Computer Video Capture



- Frame capture does not load the CPU much
- Compression demands much computation

Video Software

- Video editing and format conversion requires specialized software
- Adobe Premiere™ is an example of commercial software used for these purposes